This is the author's version of the paper Johnson, F., Sbaffi, L. & Rowley, J. *Assessing trustworthiness of digital information T*he International Data and Information Management Conference , University of Loughborough, Sept 17th 2014 which was published in the conference proceedings at https://idimc.files.wordpress.com/2015/09/idimc2014_proceedings1.pdf

Assessing trustworthiness in Digital Information

Frances Johnson, Laura Sbaffi, Jenny Rowley

Department of Languages, Information & Communications

Manchester Metropolitan University

Abstract/Summary

This paper reports the key findings of an empirical study conducted with undergraduate students to identify and assess the core constructs of online trust, with a focus on health digital information. This study suggests the use of a trust scale in further research into the potential impact of the system and information design on the users' trust and use of digital information.

1. Introduction

Within the field of Information Management there are very diverse research communities. Connecting the communities of those studying information behaviour and those focusing on systems design aims to build knowledge of information behaviours to inform and improve the design of interactive information systems. In the context of providing information in digital environments, the users' trust formation may be core as a predictor of a user's 'intention to use' a given piece of information, typically to resolve some underlying need or problem. As the design of information systems and/or information-based applications become more interactive, enhanced with functionality enabled in the web environment (for example, links to related information, recommendations, features for annotation and personalisation), it seems vital that such developments serve a purpose and, at same time, impact on the users' judgement of trust. With regards to the conference theme of 'making connections', this paper reviews our recent research on modelling online trust with a discussion on some of the main findings. Based on this, we propose a trust scale, developed for understanding the constructs of trust in digital information contexts; further research will focus on exploring the potential impact of the system design on the confident use of digital information.

Trust in digital environments has been widely studied (e.g. Chopra and Wallace, 2003; Ivanov et al., 2012; Kelton et al., 2008; Shekarpour and Katebi, 2010; Belanger and Carter, 2008; Rowley and Johnson, 2013) but, specific to online information, the user has a particular need to fulfil, creating a state of dependence on the information and providing a necessary precondition for trust to be formed (Rousseau et al., 1998). As such, trust is a dynamic concept formed in the

context of the information need. It is unlikely that use is made of information that is not trusted and so it is appropriate that, in assessing the information, we look for indicators of its trustworthiness. Within the framework of digital information, indicators of trustworthiness are likely to be numerous and the formation of trust is likely to be subject to various influences, thus becoming multidimensional and difficult to measure directly. The aim of this research is to identify the factors that influence the formation of trust in digital information. This empirical study adopted a quantitative, survey-based research design, in order to develop measurement items and explore the relationships between variables in the process of trust formation.

2. The trust scale

Trust scales have been developed in the past (e.g. the 24-item scale from Sillence et al., 2007), whilst other authors have researched trust formation by gaining insights from interviews and/or qualitative research (Robins et al., 2010; Zhang, 2012). In this study, a questionnaire was chosen to collect data as this approach was deemed the most suitable for gathering large amounts of data and collecting accurate information. The core of the questionnaire was represented by a set of 50 Likert-style statements, designed to investigate respondents' (in this case students') perceptions of the relative importance of various aspects of web health information in their evaluation of its trustworthiness. Previous research on the constructs of trust have suggested that factors like style and authority (e.g. Sillence et al., 2007) and criteria such as credibility (e.g. Corritore et al., 2012) influence the formation of trust. Research into how people evaluate information when searching in specific contexts, for example for school coursework or for health information, has identified factors relating to design as influencing trust formation (Sillence et al., 2007). Drawing on these and other previous studies, the set of statements in the questionnaire was chosen to reflect the possible constructs of trust including information credibility, usefulness, content, authority, style, verification, brand, ease of use and recommendation, all designed with a 5-point scale. For example, the variable 'authority' was indirectly measured on the responses to 5 items (e.g. 'that the author appears to be knowledgeable' and 'that the author's qualifications and/or expertise are indicated'). Each construct was represented by at least four items. The indicators of trustworthiness designed and selected for the questionnaire ranged from the tangible attributes of the information itself, such as content and style to more peripheral (but still important in the digital context) factors, such as the ease of use, as well as the factors that relate to the users' assessment of the information itself, such as usefulness and credibility. Ultimately, the implementation of the questionnaire seeks to identify which of these, as influencing factors, are the core constructs of trust formation in digital information.

Participants were 1st and 3rd year undergraduate students at a large metropolitan university in the UK recruited from the discipline areas of humanities, business, and sport. Copies of the questionnaire were distributed in class settings and participants were asked to think about some information they recently looked up on the web and which was related to some health issue that they were generally interested in or to some serious complaint/condition they had in mind. This brief scenario was set initially to ensure that all participants would be able to recall information found in response to some information need albeit at different levels of interest (i.e. passing, medium and serious). In total, 531 usable questionnaires were returned, out of 550 questionnaires originally distributed. The exploration of the data collected involved principal components analysis to test a theoretical model of the core criteria on which trust is formed and to model the observed correlations among the influences, to gain a critical evaluation of the information and its context. The approach to investigate the constructs of trust allowed for

further insight to be gained through the variations observed in the factors influencing trust formation across participants grouped by user characteristics, such as course of study or gender, as well as by task variation, such as the passing or serious interest in the health topic. Further details on these studies, including the design of the instrument alongside a detailed comparison of the items used in the evaluation across the years are reported in Rowley et al. (2014), while the factor analysis, along with modelling trust based on the criteria of usefulness and credibility and their influencing factors, is reported in Johnson et al. (2014). In this paper we consider only the difference between the two year groups of the students surveyed.

3. The constructs of trust

In order to identify the constructs of trust, principal component analysis (PCA) was used to test the convergent and discriminant validity of the questions (items) and to extract the underlying factors in the data. Subsequently, confirmatory factor analysis (CFA) was carried out on each of the 1st year and the 3rd year data sets. This was done in order to compare the resulting measures across the two year groups.

Table 1. Summary of the factors identified with CFA for the 1st and 3rd year students' datasets.

1st Year students			**3rd Year students**		
Factor	**Item**	**IR**	**Factor**	**Item**	**IR**
1 *Ease of Use – Access*	*EU1*-How easy it was to access the information	0.73	**1** *Reliable Content*	*AU4*-That the information appears to be objective (i.e. no hidden agendas)	0.65
	EU3-The information is free	0.75		*CO3*-The reliability of the information	0.73
	ST2-The ease with which I can read the information	0.77		*CO2*-The comprehensiveness of the information	0.74
	EU2-How easy it was to find the information	**0.81**		*CO4*-The accuracy of the information (such as the absence of errors)	**0.77**
2 *Believable Content*	*CR1*-Whether I feel I can believe the information	0.70	**2** *Assessing Credibility*	*CR5*-The extent to which the source contains facts rather than opinions	0.66
	AU4-The information appears to be objective (i.e. no hidden agenda)	0.73		*CR3*-The impartiality of the information	0.69
	CO4-The accuracy of the information (such as the absence of errors)	**0.75**		*CR1*-Whether I feel I can believe the information	0.70
				CR4-The quality of the information	0.75
				CR2-The objectivity of the information	**0.81**
3 *Personal Recommen.*	*RE6*-My friends and family use the source	0.71	**3** *Personal Recommen.*	*RE4*-I have seen recommendations from members of a social network community	0.71
	RE1-Family and friends have recommended the source to me	**0.86**		*RE1*-Family and friends have recommended the source to me	0.73
				RE6-My friends and family use the source	**0.79**
4 *Branded – Logo*	*BR1*-The information source features the logo of a respected brand	0.66	**4** *Ease of Use – Access*	*EU1*-How easy it was to access the information	0.89
	BR2-The information source carries the logo of a well-known brand	**0.87**		*EU2*-How easy it was to find the information	**0.97**
			5 *Assessing Usefulness*	*UF1*-That the information tells me most of what I need to know	0.78
				UF2-That the information helps me to understand the issue better	**0.88**
			6 *Style - Readable*	*ST3*-The clarity of the structure of the information	0.67
				ST1-The ease with which I can understand the information	0.85

		ST2-The ease with which I can read the information	**0.94**
	7 *Branded - Logo*	*BR1*-The information source features the logo of a respected brand	**0.90**
		BR2-The information source carries the logo of a well-known brand	**0.90**

The Cronbach's alpha coefficient was calculated; at 0.937 for the 1st year dataset and 0.933 for the 3rd year dataset, the reliability of the scale within the samples was confirmed (Bryman and Bell, 2011). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) was calculated to measure sampling adequacy and, with values of 0.879 (1st year students) and 0.874 (3rd year students), the two samples were well above the recommended value of 0.6 (Kaiser, 1974), confirming that the use principal components analysis was appropriate. A scree plot was used to identify the number of factors; also, to satisfy convergent validity, hence to make sure that all items intended to measure a construct did indeed reflect that construct, only factor loadings greater than 0.5 were selected. Items with low loading or cross loading were removed. This resulted in the identification of six factors in the 1st year dataset, explaining a total of 48% of the variance, and seven factors in the 3rd year dataset, explaining a total of 53.6% of the variance. In order to test the measurement model, CFA was performed on the factors and the items. According to Segars and Grover (1998), the measurement model should be evaluated first and then re-specified as necessary to generate the 'best-fit' model. This iterative process led to a refined measurement model with four factors and 11 items in the 1st year dataset and seven factors and 21 items in the 3rd year dataset. Item reliability (IR) ranged from 0.660 to 0.926, thus exceeding the acceptable value of 0.500 recommended by Hair et al. (2006). The average variance extracted (AVE) ranged from 0.576 to 0.624 in the 1st year and from 0.516 to 0.805 in 3rd year, which for all factors exceeded the threshold value of 0.500 recommended by Fornell and Larcker (1981). The factors from the CFA are shown in Table 1. The labels for the intended constructs were retained for each as follows: Content, Credibility, Recommendation, Ease of Use, Usefulness, Style and Brand, but for clarity each are given a sub-label reflecting the core items forming the factors extracted in the analysis.

3.1. Core constructs and impacts

The results of the principal components and confirmatory factor analyses provide the constructs involved in the formation of trustworthiness judgement of information (in health domains). The seven constructs in the 3rd year data were Assessing Credibility, Assessing Usefulness, Reliable Content, Personal Recommendation, Ease of Access, Style – Readable, Branded – Logo. These explain 53.6% of the variance in the data on the trust scales suggesting that these factors are quite comprehensive in explaining trust judgements. A smaller number of factors were found for the 1st year students, Ease of Access, Believable Content, Personal Recommendation and Branded – Logo. An in-depth examination of the differences in the factors between the 1st and 3rd years is presented in Rowley et al. (2013; 2014). The general explanation given for this difference in the principal components analysis is that the 1st year students relied more on Ease of Access while the 3rd year students demonstrated an increasing sophistication in their evaluation of the information with reference to its Credibility and Usefulness, both of which are criterion on which we might judge the information in forming trust. Based on these findings we speculate that students in their 3rd year of study engage in a more critical assessment of the information and are influenced not only by the features of the

information as indicators of trustworthiness, but also by their assessment of the information on criteria such as its Usefulness and Credibility.

Standard multiple regression was performed on both datasets to model the relationship of the factors in determining the judgements of Usefulness and Credibility. Preliminary analyses were conducted to ensure that there was no violation of the assumptions of normality, linearity, multicollinearity and homoscendasticity. The results were very different for the 1st and 3rd years; while for the 1st year no statistically significant outcome was reached, for the 3rd year the results show that the factors Ease of Access, Reliable Content and Brand associate with both Usefulness and Credibility (Figure 1a and 1b), while Style only associates with Usefulness (Figure 1b).

Figure 1. Standard multiple regression analysis performed on the 3rd year dataset posing Credibility (a) and Usefulness (b) as dependent variables.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.633	.227		7.197	.000
	Style	-.011	.058	-.012	-.194	.846
	Content	.392	.057	.419	6.912	.000
	Brand	.083	.038	.115	2.172	.031
	Ease of Access	.188	.049	.220	3.857	.000

(a) Dependent Variable: *Credibility*

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.218	.183		6.640	.000
	Style	.216	.047	.274	4.609	.000
	Content	.186	.046	.231	4.067	.000
	Brand	.112	.031	.180	3.633	.000
	Ease of Access	.141	.039	.192	3.581	.000

(b) Dependent Variable: *Usefulness*

The explanation for this in terms of a model of trust is presented in Johnson et al. (2014), who propose that users' assessment of the information on the constructs of Usefulness and Credibility are antecedents to trust formation and are determined or influenced by various factors relating to information Content, Style and Ease of Access in accessing the information. Interestingly, the factors formed in the 1st year data, Recommendation and Brand, were not found to have an association with the judgements of Usefulness and Credibility, which questions these as core constructs of trust when actively formed in assessing the information. The 1st year students appear to rely on the more contextual indicators of trustworthiness, Ease of Access, Recommendation and Brand. The assessment of the information based on the criteria of Usefulness and Credibility, with these factors relating, among others, to the Content construct, indicates a more involved and sophisticated assessment in the users' trust formation.

With the intention of extending this study of the influencing factors of trustworthiness, a subsequent study, using the same questionnaire, was carried out involving 471 adult members of the public with various educational backgrounds. The early indication in the analysis is that the data factorises closely with the 3rd year data and this will be presented along with the above findings from the earlier studies at the workshop.

4. Discussion

The findings from this study demonstrate that trust formation involves the assessment of the information based on a range of factors. This research identifies the key factors that influence trust judgements as reliability of Content, Readability/Style and Ease of Access, Brand and, to some extent, Recommendation. The factors derived from the analysis of the data positions the assessments of Credibility and Usefulness as the most important antecedents to trust. Exploration of the factors offers further insight into the critical evaluation of the information particularly with a greater range of cues and indicators being bought to bear in judgements formed by students as they progress to the final stages of their studies. The trust scale developed from the 3rd year includes the core factors Assessing Credibility and Assessing Usefulness along with the influencing factors Reliable Content, Readable Style, Ease of Access and Brand forming the 6 constructs of trust as shown in Table 2.

Table 2. The trust scale based on the 3rd year data.

Assessing Credibility The extent to which the source contains facts rather than opinions The impartiality of the information Whether I feel I can believe the information The quality of the information The objectivity of the information
Assessing Usefulness That the information tells me most of what I need to know That the information helps me to understand the issue better
Reliable Content That the information appears to be objective (i.e. no hidden agendas) The reliability of the information The comprehensiveness of the information The accuracy of the information (such as the absence of errors)
Readable Style The clarity of the structure of the information The ease with which I can understand the information The ease with which I can read the information
Ease of Access How easy it was to find the information How easy it was to access the information

Brand
The information source features the logo of a respected brand The information source carries the logo of a well-known brand

With regards to the development of our research in understanding trust, we intend to use the item scale for the identified factors to explore the influence of user and/or task characteristics on the formation of trust. Sillence et al. (2007) recognised the judgement to be dynamic, as do others such as Lucassen et al. (2013), suggesting that trust formation is further influenced by user characteristics (such as domain expertise) leading to different features of the information being used in trust judgements (Wildemuth, 2004; Hembrooke et al., 2005). Keselman et al. (2008) found that imprecise domain knowledge led consumers to search for information on irrelevant site. Whilst these investigations into the role of domain expertise focus on the impact on actions taken to find information, Lucassen et al.'s (2013) study on Wikipedia use did show that those familiar with a topic focussed on the semantic features of the information, whilst those who were unfamiliar with the topic paid more attention to surface features. Whilst the present research into trust formation did not gather details on domain knowledge or other user characteristics, a distinction was made in the 1st and 3rd year students with the 3rd year ones drawing on more and diverse factors, both relating to Content and Style as well as design factors such as Ease of Access as constructs of trust. Further research with the trust scale will explore the effect of domain knowledge and other variables, such as task and type of information, on the assessment of the information in the formation of trust judgments using the content and the design or other contextual indicators.

Further use of the trust scale is proposed with regards to the evaluation of web sites given the critical importance of trustworthiness in the provision of digital information. Participant responses to the items that formed the factors of trust provide an indication of the trust level in assessing a given piece of digital information and, in doing so, an evaluation of the impact of design. Of particular interest is the potential use of the trust scale as a diagnostic, as in the approach of the more traditional usability testing. Questionnaires used in usability studies, such as the Questionnaire for User Interaction Satisfaction (QUIS) (Chin et al., 1988) provide a series of statements for the user response to aspects of the design of the site based on usability principles and heuristics. For example, the QUIS scale measures overall reaction ratings of the system and specific factors such as interface, terminology and system feedback. This provides the designer/developer with an evaluation of the usability of the site which is an important indicator of the users' effective use and experience of the site. Based on usability principles, the questionnaire can be used as a tool to identify where the site is failing to conform to best practice and possibly hindering the overall user experience. It is proposed that the user evaluation of a web site, based on responses to the trust items, would provide a latent measure of their assessment of the information Usefulness, Credibility, Content (reliable), Style (readable), Brand and Ease of Access all of which, by influencing trust formation, indicate an evaluation of the information provided. Consequently, the trust scale could also provide a diagnostic into the impact of site design, with insights into how the information is critically evaluated. This is not to imply that trust levels might be improved simply by altering a characteristic of the information, for example in the straightforward addition of a logo. Such a step may, however, impact on the users in their assessment of the information while making trust judgments. Further research conducted within and across different information types, tasks and user groups, as suggested here, aims to confirm and develop the constructs of trust as an important tool in the evaluation of digital information and its services. The next phase of our

investigation will focus on this testing of the trust scale as an instrument for the data collection of the user evaluation of health information websites.

5. Conclusion

On the basis of these findings, recommendations for future research can be made in terms of to further developing and implementing the use of the trust scale as an instrument in advancing understanding in the information behaviour of trust formation in digital information contexts. Not least, further investigation is needed in both user and task contexts to explore the nature of the differences in the factors influencing the formation of trust with the view towards obtaining a theoretical model. On a practical application, the scale, originally developed for use in evaluating the impact of design on the user's trust formed, offers a potential to inform system designers and developers concerned with the impact on use and usability but also, and of equal importance, on the users' ability to form critical judgements on information presented and, specifically, on the particular dimensions or constructs of trust. This paper has presented a large scale study of trust in digital information with reflection on its potential to inform and build collaborative work across relevant research communities in information management. The study of trust formation, when based on the critical evaluation of the information as a key behaviour, can be invaluable in the development and evaluation of system interface design by enabling critical user behaviour.

6. References

Belanger, F. and Carter, L. (2008) Trust and risk in e-government adoption. *The Journal of Strategic Information Systems*, 17 (2): 165-176.

Bryman, A. and Nell, A. (2011) Business Research Methods. 3rd ed. Oxford University Press.

Chin, J.P., Diehl, V.A. and Norman, K.L. (1988) Development of an instrument measuring user satisfaction of the human-computer interface. In: *Proceedings of the SIGCHI conference on Human factors in computing systems* pp. 213-218, ACM.

Chopra, K. and Wallace, W.A. (2003) Trust in electronic environments. *System Sciences*, Proceedings of the 36th Annual Hawaii International Conference on IEEE.

Corritore, C., Wiedenbeck, S., Kracher, B. and Marble R.P. (2012) Online Trust and Health Information Websites. *International Journal of Technology and Human Interaction*, 8 (4): 92-115.

Fornell, C. and Larcker, D.F. (1981) Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18: 39-50.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. and Tatham, R.L. (2006) Multivariate data analysis. 6th ed. Upper Saddle River, NJ: Pearson Education.

Hembrooke, H.A., Gay, G.K. and Granka, L.A. (2005) The effects of expertise and feedback on search term selection and subsequent learning. *Journal of the American Society for Information Science*, 56 (8): 861-871.

Ivanov, I., Vajda, P., Lee, J.S. and Ebrahimi, T. (2012) In tags we trust: Trust modelling in social tagging of multimedia content. *Signal Processing Magazine*, IEEE 29 (2).

Johnson, F., Rowley, J. and Sbaffi, L. (2014) Modelling trust formation in digital information contexts. *Journal of Information Science*, in press.

Kaiser, H. (1974) An index of factorial simplicity. *Psychometrika*, 39: 401-425.

Keselman, A., Browne, A.C. and Kaufman, D.R. (2008) Consumer health information seeking as hypothesis testing. *Journal of the American Medical Informatics Association*, 15 (4): 484-495.

Kelton, K., Fleischman, K.R. and Wallace, W.A. (2008) Trust in digital information. *Journal of the American Society for Information Science and Technology*, 59 (3); 363-374.

Lucassen, T. Muilwijk, R., Noordzij, M.L. and Schraagen, J.M. (2013) Topic familiarity and information skills in online credibility evaluation. *Journal of the American Society for Information Science and Technology*, 64 (2): 254-264.

Robins, D., Holmes, J. and Stansbury, M. (2010) Consumer health information on the Web: The relationship of visual design and perceptions of credibility. *Journal of the American Society for Information Science & Technology*, 61 (1): 13-29.

Rousseau, D.M., Sitkin, S.B., Burt, R.S. and Camerer, C. (1998) Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, 23 (3): 393-404.

Rowley, J. and Johnson, F. (2013) Understanding trust formation in digital information sources: The case of Wikipedia. *Journal of Information Science*, 39 (4): 494-508.

Rowley, J., Johnson, F. and Sbaffi, L. (2013) Insights into trust in digital health information. CARPE Conference, Manchester, UK, 4-6 November 2013.

Rowley, J., Johnson, F. and Sbaffi, L. (2014) Does university education influence students approach to the evaluation of digital information? A perspective from trust in online health information. *British Journal of Educational Technology*, in press.

Segars, A.H. and Grover, V. (1998) Strategic Information Systems Planning Success: An Investigation of the Construct and Its Measurement. *MIS Quarterly*, 22 (2): 139-163.

Shekarpour, S. and Katebi, S.D. (2010) Modelling and evaluation of trust with an extension in semantic web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 8 (1): 26-36.

Sillence, E., Briggs, P., Harris, P. and Fishwick, L. (2007) How do patients evaluate and make use of online health information? *Social Science & Medicine*, 64 (9): 1853-1862.

Wildemuth, B.M. (2004) The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science and Technology*, 55 (3): 246-258.

Zhang, Y. (2012) College students' uses and perceptions of social networking sites for health and wellness information. *Information Research – an International Electronic Journal*, 17 (3).